確率密度関数の定義再訪

石川 徹也　<tiskw111@gmail.com>

2022 年 6 月 26 日

はじめに

現代の確率論は，その歴史的過程により，測度論を用いて高度に形式化されています．形式化されること自体は，厳密さを尊ぶ現代数学の理念に叶うものであり，大いに歓迎されるべきなのですが，その一方で，残念なことに，過度な形式化は概念の本質を覆い隠してしまうことがあります．その最たるもののひとつが，確率密度関数です．

今，試みに確率の書籍を開いて確率密度関数の定義を見てみましょう．伊藤 [1] には次のように定義されています．

定理 0.1

確率密度関数の定義 確率空間 $(\Omega, \mathcal{F}, P)$ を考える．*Radon – Nikodym* の定理より，任意の事象 $E \in \mathcal{F}$ に対し

$$P(E) = \int_E p(\omega) \mathrm{d}\mu(\omega), \tag{1}$$

をみたす関数 $p(\omega)$ が存在する（*Radon – Nikodym* 微分）．これを確率測度 P の確率密度と呼ぶ．これは微分形

$$p(\omega) \triangleq \frac{\mathrm{d}P}{\mathrm{d}\mu}(\omega), \tag{2}$$

で表されることもある．

上記は任意の確率空間に対して適用できるよう抽象化された表現であり，さすがにちょっと分かりにくいので，実数体に値を取る確率変数 X に限定した表現に直してみましょう．詳細は省略しますが，実数値確率変数 X の場合，上記の確率密度関数は

$$p_X(x) \triangleq \frac{\mathrm{d}P(X \leq x)}{\mathrm{d}x}, \tag{3}$$

と書くことができます．累積分布関数 $f_X(x) \triangleq P(X \leq x)$ を導入すれば

$$p_X(x) \triangleq \frac{\mathrm{d}f_X}{\mathrm{d}x}, \tag{4}$$

と表現するのと同じことです（何となく抽象的な定義の微分形に似ていることが分かるかと思います）．

ある程度確率論に慣れた経験者にとって，式 (3) や 式 (4) は「なるほど上手く定義したものだ」と思える代物だと思います．ですが，そもそも確率密度関数の意味を知らない確率論の初学者に式 (3) や 式 (4) を見せたところで

「で，確率密度関数って結局何なの？」

という疑問が浮かぶだけでありましょう．このような疑問が浮かぶのも至極当然です．式 (3) や 式 (4) は，形式的に数学を構築していくには大変都合の良い定義ですが，確率密度関数とは何かという本質的な問いには全く答えていない式だからです．これが過度な形式化の弊害なのです．

本文書の目的は「確率密度関数とは何か？」という疑問に対し，定義式を見せる以外の手段で回答を試みることにあります．もし本文書を読み終えて式 (3) を見直したとき，少しでも「なるほど上手い定義式だ」と感じて頂けたのであれば，本文書は成功だと言えるでしょう．

1　古典的確率論の破綻

そもそも，なぜ確率密度関数が必要なのだろうか．実は目的は非常にシンプルであり「ある確率変数 X が特定の値 x をとる確率を表現したい」というだけである．こんな簡単なことが，実は連続値を取る確率変数では表現が難しいのである．

有限個の値しか取らない確率変数ではこのような困難は生じない．まずはそれを確認するところから始めよう．有限個の値しか取らない確率変数として，正しく作られたサイコロの出る目を表す確率変数 D を考えてみよう．確率変数 D は $1, 2, 3, 4, 5, 6$ の 6 通りの値を取り，その確率はそれぞれ

$$P(D = 1) = P(D = 2) = \cdots = P(D = 6) = \frac{1}{6}, \tag{5}$$

である．確率変数 D がどの値をとるにしろ，その確率は上記の通り簡単に表現ができる．何の問題もない．

しかしこれが連続値を取る確率変数では簡単ではない．連続値を取る確率変数として，ここでは区間 $[0, 1]$ 上で一様分布にしたがう実数値の確率変数 X を考えてみよう．区間 $[0, 1]$ には無限個の実数が含まれていることに注意されたい．さて，確率変数 X は一様分布にしたがうのであるから，

$$P(X = 0) = \cdots = P(X = 0.5) = \cdots = P(X = 1), \tag{6}$$

である．しかし確率変数 X の取りうる値は無限個あるのに対し，確率の総和は 1 であるから，

$$P(X = 0) = \cdots = P(X = 0.5) = \cdots = P(X = 1) = 0, \tag{7}$$

とならざるをえない（この確率が 0 になってしまうことの厳密な説明は次章に行うので，今は直感的に理解ができれば十分である）．しかしこれは明らかに表現として不十分である．式 (7) は決して間違えてはいない．しかし式 (7) を見るとあたかも確

率変数はどの値も取らないように見えてしまう．しかし実際にそんなことはなく，確率変数 X は区間 $[0, 1]$ 上の一様分布なのである．

つまり，連続値を取る確率変数 X に対しては，$X = x$ となる確率を定義するのに通常の $P(X = x)$ という表現は，結果が常に 0 になってしまうので，適さないのである．ではどうすれば良いのだろうか．

この困難を解決してくれるのが確率密度関数なのである．次章以降でこれを説明していこう．

2　Reimann 積分論的確率論

TBD

3　確率密度関数

TBD

4　累積分布関数との関係

TBD

おわりに

TBD

補足 A：記号一覧

本節では，本文書で用いられている数学記号を簡潔にまとめておく．

参考文献

[1] 伊藤清，「確率論の基礎」，岩波書店，2004.

記号	意味	使用例
$\triangleq$	記号 $\triangleq$ の左辺を右辺で定義する．等号 $=$ に近い意味合いを持つが，変数や関数の定義を行う際などに，それが「定義だ」ということを明示したいときに使用する．記号 $:=$ を用いることもある．	$f(x) \triangleq x^2$
$\frac{\mathrm{d}f}{\mathrm{d}x}$	関数 f の変数 x による微分．	$\frac{\mathrm{d}f}{\mathrm{d}x} = 2x$
$\frac{\partial f}{\partial x}$	関数 f の変数 x による偏微分．	$\frac{\partial f}{\partial x} = 2x$

T. Ishikawa